R AS THE ENVIRONMENT FOR DATA ANALYSIS IN PSYCHOLOGICAL EVALUATION

Ana María Ruiz-Ruano and Jorge L. Puga
UCAM Universidad Católica de Murcia

R es un entorno libre para el análisis estadístico de datos y para la generación de gráficos que está cobrando un papel clave en un amplio espectro de áreas de conocimiento. Pese a que el usuario ha de interaccionar con el programa, principalmente, por medio de una consola de comandos; se están invirtiendo esfuerzos en proporcionar entornos gráficos amigables que aproximen al usuario novel a esta especie de lengua franca del análisis estadístico contemporáneo. Además, la flexibilidad y versatilidad del programa, en conjunción con una comunidad de desarrolladores y usuarios global, facilitan que R sea la herramienta elegida en muchos contextos básicos o aplicados. En este trabajo se presenta la utilidad potencial que supone R para la evaluación psicológica en general a la vez que se acentúa el rol que asume la estadística propiamente dicha como instrumento de progreso de la psicología al amparo del método científico.
Palabras clave: R, Evaluación psicológica, Análisis estadístico, Psicometría.

R is a free computing environment for statistical data analysis and graph creation. It is becoming a key tool in a wide range of knowledge domains. The interaction with the software is mainly based on a command line interface but efforts are currently being made to develop friendlier graphical user interfaces that will help novice users to become familiar with this programming language. R is a flexible and powerful system thanks to the development community that is working together to improve its capabilities. As a result, it is the chosen statistical software in many applied and basic contexts. This paper highlights the potential usefulness of R for psychological assessment and related areas. Additionally, the relevance of statistical data analysis is emphasised as an instrument that will boost the progress of psychology under the umbrella of the scientific method.
Key words: R, Psychological assessment, Statistical analysis, Psychometrics.

T he statistical analysis of data has become fundamental in contemporary culture. It is therefore not surprising that some degree of statistical literacy is required in society as a whole in order to aspire to new heights of human development (e.g., Bond, 2009; Seldmeier & Gigerenzer, 2001). For science, which seeks to understand nature, and for technology that yearns to control nature, statistics is fading as the key tool that guides and leads the progress of human communities. However, although today we have huge computational and statistical algorithms and we can implement them in computers powerful enough to handle a considerably large volume of data, it appears that we are not able to take advantage of this situation. It is worth remembering the slogan popularized by Pirelli (the Italian tyre company) in the mid-1990s, because it could clearly be applied to this situation: "power is nothing without control". Today we have very powerful computers, computer software for data analysis that is easy to use, and it is also relatively easy to collect data to carry out studies. However, in spite of all these facilities (power), it is crucial to have a good plan of what you want to do and a good research design (control), taking into account the data analysis to be performed *a posteriori*, to test the initial hypothesis (e.g., Bakan, 1966; Cohen, 1990; Dar, Serlin & Omer, 1994; Munafò et al, 2014; Rosnow & Rosenthal, 1989). In other words, even though we may have all the tools and all the data at our disposal, they will not do us any good if our study is lacking in planning and design. And, contrary to what one might think, being flexible in

data analysis that is performed after an abnormal or non-existent planning phase does not help the investigation to progress satisfactorily, but rather leads to fruitless or contrived results with no scientific meaning (Ioannidis, 2005).

In recent years, there have been numerous reports of malpractice in research (e.g., Button et al, 2013; Cohen, 1990; Gigerenzer, 2004; Munafò et al, 2014; Trafimow & Marks, 2015; Vaux, 2012; Weissgerber, Milic, Winham & Garovic, 2015) which present a panorama that is scary, to say the least, with regards to the analysis of statistical data. There are even authors like Ioannidis (2005) who dare to affirm that many of the research papers presented are false and that, in fact, most of the discoveries that are made are merely evidence of these biases existing in the various sub-fields of research. As for malpractice, for example, Munafò et al. (2014) criticise the quality systems prevailing in science, the low reproducibility of the studies, the existence of a bias in scientific publications sponsored by the editorial policies of the scientific journals, the falsifying of the data, the low statistical power of the results presented, the development of illicit practices in research and the *p* value used widely as a critical element of statistical inference. Meanwhile, Button et al. (2013) focus on denouncing the low statistical power presented in the results of neuroscience research and the small sample size used in most studies. The importance of taking into consideration the ethical implications of animal sacrifice is also highlighted (see also e.g., Cressey, 2015, Nature Publishing, 2015) as well as the financial investment that is being made only to fail to obtain, in the majority of cases, useful results for society.

One of the issues that continues to receive most attention in terms of its misuse is the null hypothesis significance testing procedure or NHSTP (e.g., Bakan, 1966; Cohen, 1994; Gigerenzer, 1998; Haller & Krauss,

2002; Halsey, Curran-Everett, Vowler & Drummond, 2015; Nieuwenhuis, Forstmann & Wagenmakers, 2011; Rosnow & Rosenthal, 1989; Trafimow, 2014; Trafimow & Marks, 2015) and the abusive use of the p value as a single element that guides statistical inference in scientific studies. These malpractices were highlighted long ago by Bakan (1966), who referred to it as a flagrant attack against common logic and affirmed that "when we reach a point where statistical procedures are substitutes rather than aids to thinking, and we reach absurd conclusions, then we must return to the basics of common sense" (p.436). Similarly, Cohen (1994) and Gigerenzer (1998) warned of the importance of making good use of hypothesis testing and not turning it into a ritual that has been institutionalised in a scientific culture that has systematically repeated the same mistakes throughout history (Gigerenzer, 2004).

One of the most drastic and controversial steps that have been taken recently regarding the use of testing the null hypotheses was from the 2014 editorial of the journal *Basic and Applied Social Psychology* (Trafimow, 2014), in which it was proposed that authors should use a different statistical methodology than the null hypothesis procedure that used the p value as the sole value to guide the statistical inference, and which was explicitly prohibited in the 2015 editorial (Trafimow & Marks, 2015). The authors are asked to exploit the data at the descriptive level, and to use another type of data analysis that might lead to higher quality scientific results. The editors point out that, contrary to what many researchers may think, $p <.05$ is very easy to obtain and sometimes even serves as an excuse to publish poor quality research studies. There is evidence to prove the misuse of the p value in a statistical test. For example, the study carried out by Haller and Krauss (2002) who, after observing the misinterpretations made of the p value in psychology, concluded that both teachers and students had deficiencies when interpreting the significance of the research results. However, Leek and Peng (2015) suggest that the aforementioned ban will not guarantee an increase in the quality of the *science* that is produced, as well as the fact that the p value is "the tip of the iceberg" in the world of statistics and data analysis.

Over the years, psychology as a science has provided a number of measurement tools to carry out work in different areas of study that are closely linked to statistical data analysis. In neuropsychology, clinical, educational and organisational psychology, for example, it is very common to use tests. In fact, as noted by Hernández, Tomás, Ferreres & Lloret (2015), tests are considered basic instruments in psychological assessment to guide the psychologist's decision making. This is why, in order to be used correctly, tests must meet a series of psychometric guarantees, which should be available to the user applying the test. Hernández et al. (2015) note that some of the information that is not provided and that would be very useful is information related to the evidence of content validity, predictive validity, reliability as stability in the measurement, analysis of bias or differential item functioning. They also point out that it is the responsibility of the professional, even if their field of work may not be psychometrics, to train and learn about psychometric advances. However, away from the instruments used in psychological intervention or research, psychology as a science needs the scientific method in daily practice. Statistical data analysis plays an important role in all of these situations, which are not foreign to the professional psychologist. Therefore, like Dar et al. (1994), we could argue that the key to performing good data analysis, regardless of the area of psychological knowledge we are considering, lies in starting from a good base theory and designing impeccably the research you wish to carry out in order to prevent the weaknesses that are attributed to the procedure of testing null hypotheses.

The main objective of this paper is to present the R software as an environment for performing statistical analysis that may be very useful for psychology in general. As we will see below, this environment can enhance the learning of the statistical technique we are using. The next section is about the environment R, which is presented from a historical and philosophical perspective, as well as the perspective of the user interaction with R. Next, we briefly review some of the uses that R offers for psychological evaluation, presenting the different packages and their applicability to this discipline. Finally, we end with a reflection on possible solutions and improvements for attempting to implement useful and quality research for society in general with the support of statistical analysis, in which R plays a crucial role today.

R ENVIRONMENT

R is a working environment for carrying out statistical analysis and creating graphs (R Development Core Team, 2011). The graphical interface of the program is a command console, which means that in order to interact with it you have to write lines of code and run them. We can trace the origin of R back to the 1970s, when the term S was coined to refer to a high-level programming language designed to perform statistical computations. The shift to R occurred in the 1990s, coinciding with the massive expansion of S-Plus, the commercial version of S. Ross Ihaka and Robert Gentleman (University of Auckland, New Zealand) wrote an educational version of the first distribution of R, which was released to the general scientific community in 1996 (Ihaka & Gentleman, 1996). R was structured based on the programming language S and Scheme, a dialect of Lisp. From the beginning, the aim was for it to be a multiplatform language that could be used under different operating systems. In 1997 the R Core Team was created. This was the body responsible for developing the software base and providing continuous support for the development and distribution of R (Fox, 2009). Since then, while continuing to grow and develop, R has established itself as the statistical reference tool par excellence. In fact, Tippmann (2015) notes that R is the data analysis software most used in 2015 on a par with, or even surpassing, other commercial software programs.

The fact that it is considered a working environment means that R is a programming language that comes equipped with a set of tools for calculating and generating statistical graphics (Ihaka & Gentleman, 1996). This working environment was made to be –and still is– free, in the sense that it is not necessary to pay to use, copy or distribute the program (Carleos & Corral, 2013). Furthermore, it is protected against possible privatisation / commercialisation under a GNU license. The GNU GPL (General Public License) philosophy is the basis of freedom and openness of code, and its purpose is to protect the rights and freedoms of end users (http://www.fsf.org). The basic pillars on which R rests are allowing to use, share, study and modify the software code that is at the base of the program. This means that any user of R can develop a new application of it, modifying what already exists, sharing it and using it freely.

R consists of a *base* structure that contains a number of basic functions for making statistical calculations and graphs. Additionally, a range of

accessories called packages have been created by the scientific community to perform statistical analyses of more specific data. This software development activity may be regarded as having been done in an altruistic way and caters more to the intrinsic factors of the person – such as the satisfaction produced by this collaboration, the intrinsic rewards of the job or the fact that they are contributing to a public asset– than to external factors such as scientific recognition or money (Fox, 2009). An R package is simply "a set of functions that maintain some kind of relationship between them" (Elosua, 2011, pp.24-25). For example, there are packages for the statistical analysis of issues as varied as the content of electronic texts, the spatial distribution of urban areas or the neural activation recorded by a neuroimaging technique. All of these are freely available on the Internet at no cost. The CRAN (Comprehensive R Archive Network, http://www.r-project.org) is the network location where all of the downloads are managed. For greater detail on downloading and installing R and its packages see, for example, Elosua (2009) or López (2012, 2013).

The main advantage of R compared with other statistical analysis software is the freedom and the fact that it is free to use. Other advantages of using this working environment include the fact that it promotes the learning of basic statistics (López, 2012; Tippmann, 2015) and that it can be run on different operating systems (Fox, 2009), such as Windows, Mac OS or Linux. Furthermore, Huber et al. (2015) note that the use of R as a statistical program facilitates research and innovation, because it allows the development of rapid prototypes; it is flexible and functional; it allows reproducibility; changes or modifications are made quickly; it provides graphic facilities and it permits interaction with other programming languages such as C and C ++ or JavaScript for web applications. With regard to the interpretation of the outputs generated after performing statistical analysis, certain advantages can also be observed. Since the user of the software has to become more involved in the statistical tool he is using from a technical point of view, he ends up being more adept at interpreting the results generated by the data analysis he applies. One drawback we can point out is the difficulty working with a command environment compared to other programs or interfaces where all you have to do is click (e.g., López, 2013; Tippmann, 2015). However, as discussed below, the graphical user interfaces that have been developed are increasing in number and in quality in order to make the interaction with R more user friendly.

The graphic interface R, as we mentioned above, is based on command lines. However, to facilitate interaction with this program, graphical interfaces have been developed that are more intuitive for the users. These interfaces can be classified into two types, the first are Windows type menus (e.g., R Commander or RKWard) and the second are code editors (e.g., Tinn-R, Emacs or RStudio). Interfaces with Windows-type menus are preferred by users that are familiar with commercial software (e.g., Elosua, 2009; Fox, 2005; López, 2013). The R Commander package developed by Fox (2005) was presented as the most natural transition between commercial statistical analysis software and R (Elosua, 2009). It was developed with the objective of providing a user-friendly platform for basic courses in statistics, aiming to reduce the possibility of committing careless mistakes (e.g., calculating the mean of a nominal variable) and, finally, to make visible the relationship established between what is selected in the menus and the R code that would have had to be used in the command console. Fox (2005) points

out that this type of graphical interface has positive and negative aspects for the users. One positive point noted is that it is not necessary to remember commands or arguments of the functions, which reduces the likelihood of errors in the script syntax, and also it is a comfortable environment for the work of novice, infrequent or casual users. The negative aspects highlighted were the fact that it is tedious and difficult to reproduce certain statistical analyses, because it involves navigating through multiple dialogue boxes, and also the fact of having to incorporate numerous analyses on the interface, which can end up becoming a labyrinth that is difficult to use (Fox, 2005). One of the most attractive proposals that have been made in recent times to facilitate the user interaction with R base code can be found in the program JASP (Love et al., 2015). JASP (https://jasp-stats.org) aims to enrich the user experience with a quick, agile and easy to use graphical interface. Meanwhile, interfaces such as RStudio (www.rstudio.com), which is an integrated or interactive development environment, –meaning it is an application that provides facilities and tools to users that work in computer programming– are preferred by users with more advanced programming skills in R or those who wish to try to be productive with code (López, 2013).

Finally, we would like to note that turning to the sources of support for working with R is almost inevitable regardless of the level of skill we have in using it. R comes with a series of PDF manuals that are available to the user from the console itself. It also provides various other sources of support (in html, FAQs, etc.). Another way of obtaining help, in this case regarding specific packages, is by going directly to the directory where the package is hosted on the web or to the local installation directory. Using traditional web browsers can be useful, the Quick-R (www.statmethods.net) or the CRAN webpage (López, 2013). There are manuals in both English (e.g., Field, Miles & Field, 2012) and Spanish (e.g., Arriaza et al, 2008; Elosua, 2011; Elosua & Etxeberria, 2012; López, 2012, 2013) to facilitate the interaction with this software. Therefore, in addition to being supported by a community of developers and users of the software, the researchers and professionals who prefer to use R as a tool for the statistical analysis of data will have a large pool of resources to make their interaction with the software more satisfactory. Table 1 presents a short list of resources, freely available on the web, which can be useful to make the interaction with R more user-friendly.

USES IN PSYCHOLOGICAL ASSESSMENT

In this section we present very briefly some of the R packages that take centre stage in the field of psychological assessment and particularly we discuss some of those most related to psychometric technology. It would be impractical to try to cover them all and that is not the intention of this article. At the moment of writing these words, there are 6,695 packages available for download and it is likely that by the time the article is being read this amount will have increased by a substantial number given the quasi-exponential growth experienced in recent years (Elosua & Etxeberria, 2012). Although only some of the packages are directly related to psychology, it is also true that there are many packages that may be used by our discipline occasionally, even though they have been developed in areas as diverse as topographic pattern recognition. In any case, CRAN contains a listing of the available packages together with documentation that specifies how they can be used efficiently.

In 2007 Patrick Mair and Reinhold Hatzinger inaugurated the *Psychometric Task View* with the aim of bringing together all the initiatives and progress that had been made so far in the field of psychometrics (Mair & Hatzinger, 2007a). The same year, a special issue appeared in the *Journal of Statistical Software* which formally presented some of the most relevant packages for psychometric data analysis with R (de Leeuw & Mair, 2007). The packages (127 at present) were grouped into five major groups (plus one miscellaneous) oriented to psychometric modelling under item response theory, correspondence analysis, developing structural equation models, multidimensional scaling, and classical test theory.

Most of the packages compiled by the *Task View* are concentrated in the group dedicated to statistical modelling under item response theory (IRT). Thus it could be deemed to fit in with the attempt being made in recent years to advance the advantages offered by this theory of psychological measurement compared to Classical Test Theory (Muñiz, 1997; 2010). If we had to highlight any of the packages included in this group, it would be **ltm** (Rizopoulos, 2006) since as well as including the functionality to develop dichotomous and polytomous IRT models (one, two or three parameters), it also contains functions that have been highlighted in different application contexts other than IRT (Falissard, 2012). We should also highlight the packages **eRm** (Mair & Hatzinger, 2007b) and **plRasch** (Anderson, Li, & Vermunt, 2007).

Additionally, we should emphasise four essential packages for working from the perspective of Classical Test Theory (Mair & Hatzinger, 2007a): **psychometric**, **psy**, **psych** and **MiscPsycho**. These four packages bring together basic and advanced functions dedicated to item analysis, and the study of the validity and reliability of tests, as well as useful functions for the development of scales under classical philosophy.

Another large group of packages is the one intended for factor analysis, principal component analysis and the development of structural equation modelling. Factor analysis can be considered a statistical analysis technique that, despite having emerged within psychology, shows its usefulness in knowledge areas as diverse as biology or economics (Ferrando & Anguiano-Carrasco, 2010). Classical factor analysis and principal component analysis are implemented in the **stats** package that is contained in the base distribution of R. Although the versatility of the factanal() and princomp() functions of the **stats** package enables you to run factor analysis and basic principal component analysis, other packages have been developed to complement and optimise the user experience of R in this regard. Structural Equation Models are emerging today as a kind of evolution of factor analysis and associated techniques although they involve even greater complexity from the point of view of statistical estimation (Ruiz, Pardo & San Martin, 2010). In R there are some packages that have been aimed at developing this type of model with similar functionality to the commercial packages available to date. For example, the **sem** package (Fox, 2006) and the **lavaan** package (Rosseel, 2012) are two tools that are hugely popular for the development of structural equation models that allow the estimation of factor models of various kinds while also facilitating their graphical representation. Packages such as **polycor**, among others, for estimating models involving polychoric correlations, **systemfit**, for nonlinear structural alternatives, and **pls**, for estimating the minimum partial quadratic allow great flexibility and versatility in the range of statistical techniques applicable to data generated in the field of study of psychology that can be treated under the structural perspective.

Finally, we should note that different procedures aimed at estimating and testing models of multidimensional scaling can be found in packages such as **MASS**, **MLDS**, **vegan**, **labdsv**, **ecodist** and **ade4**.

CONCLUSIONS

As we have tried to demonstrate, R is a considerably versatile and flexible tool that enables us to carry out statistical data analysis that is difficult to achieve using any other software available today. In addition, the large number of packages that are associated with R makes it easy for it to be used as the only software, which avoids the need to switch from one program to another when specific computing needs arise. Thus, psychology could benefit from this potential in all of its knowledge areas. However, it is also true that interacting efficiently with the program requires a certain amount of skill, which cannot be acquired quickly. In any case, more and more attempts are being made to develop graphical user interfaces that are easier for general users, so that they do not panic when faced with the command console. Let us not forget, as we have pointed out from the beginning, that *power without control is useless* and that statistical data analysis only provides a spectrum of techniques to be applied which have to be backed up by solid scientific theories.

Cohen, (1990) noted that statistical inference, in conjunction with the informed judgment of scientists is a very useful tool. However, he also

---

TABLE 1
SOME FREE RESOURCES FROM THE WEB THAT CAN BE USEFUL TO MAKE THE INTERACTION WITH R MORE USER-FRIENDLY

Resource Description	Link
Brief manual covering basic and common statistical analysis in Psychology with R Commander (Spanish)	http://hdl.handle.net/10835/1658
One of the best compilations of tutorials on R, which addresses both the basic and more advanced aspects of statistical analysis for Psychology (English)	http://ww2.coastal.edu/kingw/statistics/R-tutorials/
This page contains a set of links to introductory video tutorials on R (Spanish)	http://www.jpuga.es/Docencia/mooc-r.html
Interactive course on operating the R console (English)	http://tryr.codeschool.com/
Basic manual on psychometrics with R Commander (Spanish)	https://web-argitalpena.adm.ehu.es/pasa_pdf.asp?File=UWLGPS5661

believed that statistics are not the most important tool of science but rather the result of what comes before (theory, planning, scenario, etc.). This is why we need to re-emphasise the importance of research planning as an essential part of the scientific method, as many other authors also affirm (e.g., Cohen, 1990; Munafò et al, 2014), so we can obtain results that really are useful for both the scientific community and society in general.

In this sense, a cultural change in the scientific community is necessary (e.g., Morrison, 2012; Munafò et al, 2014; Weissgerber et al, 2015) to allow us to obtain quality work, faithful to professional ethics adapted to the current times, along with a change in the editorial policies of the scientific journals (Weissgerber et al., 2015). This cultural change will improve in part if researchers receive training (Weissgerber et al., 2015), as Leek and Peng (2015) point out when they say that "education is the beginning." As for the changes in editorial policies, at least in psychology, we have the somewhat extreme example of the *Basic and Applied Social Psychology* journal which prohibited authors from sending articles that use the *p* value (Trafimow, 2014; Trafimow & Marks, 2015).

Similarly, and to provide greater quality to the scientific results that are presented, some authors recommend that the exploratory data should be emphasised, i.e., that greater use should be made of descriptive statistics (e.g., Cohen, 1990, 1994; Leek & Peng, 2015; Trafimow & Marks, 2015; Weissgerber et al, 2015). It would also be interesting to provide more graphical results (Cohen, 1994), information on the effect size (e.g., Cohen, 1994), confidence intervals or using Bayesian statistics (e.g., Bakan, 1966; Cohen, 1992, 1994; Haller & Krauss, 2002; Puga, Krzywinski & Altman, 2015a, 2015b; Trafimow & Marks, 2015). Another important point suggested by some authors is to carry out further replications of studies that find results that may be considered weak from a methodological point of view (e.g., Halsey et al. 2015; Huber et al, 2015; Munafò et al, 2014; Rosnow & Rosenthal, 1989). In addition, Ioannidis (2005) suggests that obtaining meaningful results should not be pursued and the previous information regarding the issue being researched should be taken into account. The latter underscores the interest in using Bayesian statistics, the foundations of which are based on pre-existing information.

The credibility of the science of psychology is in our hands and the statistical analysis of data is an essential ally in which to trust in order to optimise the evolution of scientific progress. R is presented today as the lingua franca of statistical analysis and graph generation in a wide range of subject areas. It is about time it was determined whether R will ultimately become something that could be considered as a world heritage (López, 2012) or whether its peak will simply be the result of a passing trend.

ACKNOWLEDGEMENTS

REFERENCES

Arriaza, A. J., Fernández, F., López, M. A., Muñoz, M., Pérez, S., & Sánchez, A. (2008). *Estadística básica con R y R-Commander*. [*Basic statistics with R and R-Commander*.] Cádiz: Servicio de Publicaciones de la Universidad de Cádiz.

Anderson, C. J., Li, Z., & Vermunt, J. K. (2007). Estimation of models in a Rasch family for polytomous items and multiple latent variables. *Journal of Statistical Software, 20*, 1-36.

Bakan, D. (1966). The test of significance in psychological research. *Psychological Bulletin, 66*, 423-437. doi: 10.1037/h0020412

Bond, M. (2009, October). Decision-making: risk school. *Nature, 461*, 1189–1192. doi: 10.1038/4611189a

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Neuroscience, 14*, 365-176. doi: 10.1038/nrn3475

Carleos, C. & Corral, N. (2013). Paquetes estadísticos con licencia libre (I). [Statistical packages with free licences (I)] *Revista Electrónica de Metodología Aplicada, 18*, 12-37.

Cohen, J. (1990). Things I have learned (so far). *American Psychologist, 45*, 1304-1312. doi: 10.1037/0003-066X.45.12.1304

Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science, 1*(3), 98-101. doi: 10.1111/1467-8721.ep10768783

Cohen, J. (1994). The earth is round (*p* < .05). *American Psychologist, 49*, 997-1003. doi: 10.1037/0003-066X.50.12.1103

Cressey, D. (2015, Abril). UK funders demand strong statistics for animal studies. *Nature, 520*, 271-272. doi:10.1038/520271a

Dar, R., Serlin, R. C., & Omer, H. (1994). Misuse of statistical tests in three decades of psychotherapy research. *Journal of Consulting and Clinical Psychology, 62*, 75-82. doi: 10.1037/0022-006X.62.1.75

de Leeuw, J. & Mair, P. (2007). An introduction to the special volume on "psychometrics in R". *Journal of Statistical Software, 20*, 1-5.

Nature Editorial (2015, April). Numbers matter. *Nature, 520*, 263-264. doi:10.1038/520263b

Elosua, P. (2009). ¿Existe vida más allá de SPSS? Descubre R. [Is there life after SPSS? Introducing R] *Psicothema, 21*, 652-655.

Elosua, P. (2011). *Introducción al entorno R.* [*An introduction to the R environment*] Bilbao: Euskal Herriko Univertsitateko Argitalpen Zerbitzua / Editorial Service of the University of the Basque Country.

Elosua, P. & Etxeberria, J. (2012). *R Commander. Gestión y análisis de datos* [*R Commander. Data management and analysis*]. Madrid: La Muralla.

Falissard, B. (2012). *Analysis of questionnaire data with R*. Boca Raton, FL: Taylor & Francis Group.

Ferrando, P. J. & Anguiano-Carrasco, C. (2010). Factor analysis as a research technique in psychology. *Papeles del Psicólogo, 31*, 18-33.

Field, A., Miles, J., R Commander. Gestión y análisis de datos Field, Z. (2012). *Discovering statistics using R*. London: SAGE.

Fox, J. (2005). The R commander. A basic-statistics graphical user interface to R. *Journal of Statistical Software, 14*, 1-42.

Fox, J. (2006). Structural equation modeling with the **sem** package in R. *Structural Equation Modeling, 13*, 465-486.

Fox, J. (2009). Aspects of the social organization and trajectory of the R project. *The R Journal, 1/2*, 5-13.

Gigerenzer, G. (2004). Mindless statistics. *The Journal of Socio-Economics, 33*, 587-606. doi: 10.1016/j.socec.2004.09.033

Gigerenzer, G. (1998). We need statistical thinking, not statistical rituals. *Behavioral and Brain Sciences, 21*, 199-200. doi: 10.1017/S0140525X98281167

Haller, H. & Krauss, S. (2002). Misinterpretations of significance: A problem students share with their teachers? *Methods of Psychological Research Online, 7*, 1-20.

Hasley, L. G., Curran-Everett, D., Vowler, S. I., & Drummond, G. B. (2015). The fickle *P* value generates irreproducible results. *Nature Methods*, *12*, 179-185. doi: 10.1038/nmeth.3288

Hernández, A., Tomás, I., Ferreres, A., & Lloret, S. (2015). Tercera evaluación de tests editados en España [Third evaluation of tests published in Spain]. *Papeles del Psicólogo*, *36*, 1-8.

Huber, W., Carey, V. J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B. S., … & Morgan, M. (2015). Orchestrating high-throughput genomic analysis with Bioconductor. *Nature Methods*, *12*, 115-121. doi: 10.1038/nmeth.3252

Ihaka, R. & Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, *5*, 299-314. doi: 10.1080/10618600.1996.10474713

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLOS Medicine*, *2*, 696-701. doi: 10.1371/journal.pmed.0020124

Leek, J. T. & Peng, R. D. (2015, Abril). P values are just the tip of the iceberg, *Nature*, *520*, 612. doi: 10.1038/520612a

López, J. (2012). *Introducción al análisis de datos con R y R Commander en psicología y educación [Introduction to data analysis with R and R Commander in psychology and education.]*. Almería: Editorial of the University of Almería. Available from http://hdl.handle.net/10835/1658.

López, J. (2013). *Primeros pasos analizando datos estadísticos con R. Un texto para acompañarte [First steps to analysing statistical data with R. A text to accompany you.]*. Murcia: Universidad Católica San Antonio.

Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, A. J., Ly, A., Gronau, Q. F., Smira, M., Epskamp, S., Matzke, D., Wild, A., Rouder, J. N., Morey, R. D. & Wagenmakers, E.-J. (2015). JASP (Version 0.7)[IT program].

Mair, P. & Hatzinger, R. (2007a). Psychometrics task view. *R News*, *7*, 38-40.

Mair, P. & Hatzinger, R. (2007b). Extended rasch modeling: The eRm package for the application of IRT models in R. *Journal of Statistical Software*, *20*, 1-20.

Morrison, R. P. (2012). Reforming science: methodological and cultural reforms. *Infection and Immunity*, *80*, 891-896. doi: 10.1128/IAI.06183-11

Munafò, M., Noble, S., Browne, W. J., Brunner, D., Button, K., Ferreira, J., … & Blumenstein, R. (2014). Scientific rigor and the art of motorcycle maintenance. *Nature Biotechnology*, *32*, 871-873. doi:10.1038/nbt.3004

Muñiz, J. (1997). *Introducción a la Teoría de Respuesta a los Ítems [Introduction to Item Response Theory]*. Madrid: Pirámide.

Muñiz, J. (2010). La teoría de los tests: Teoría Clásica y Teoría de Respuesta a los Ítems. [Test Theory: Classical Theory and Item Response Theory] *Papeles del Psicólogo*, *31*, 57-66.

Nieuwenhuis, S., Forstmann, B. U., & Wagenmakers, E. (2011). Erroneous analyses of interactions in neuroscience: a problem of significance. *Nature Neuroscience*, *14*, 1105-1107. doi: 10.1038/nn.2886

Puga, J. L., Krzywinski, M., & Altman, N. (2015a). Points of Significance: Bayes' theorem. *Nature Methods*, *12*, 277–278. doi:10.1038/nmeth.3335

Puga, J. L., Krzywinski, M., & Altman, N. (2015b). Points of Significance: Bayesian statistics. *Nature Methods*, *12*, 377–378. doi: 10.1038/nmeth.3368

R Development Core Team. (2011). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rizopoulos, D. (2006). ltm: An R package for latent variable modeling and item response theory analyses. *Journal of Statistical Software*, *17*, 1-25.

Rosnow, R. L. & Rosenthal, R. (1989). Statistical procedures and the justification of knowledge in psychological science. *American Psychologist*, *44*, 1276-1284. doi: 10.1037/0003-066X.44.10.1276

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*, 1-36.

Ruiz, M. A., Pardo, A., & San Martín, R. (2010). Structural equation models. *Papeles del Psicólogo*, *31*, 34-45.

Seldmeier, P. & Gigerenzer, G. (2001). Teaching bayesian reasoning in less than two hours. *Journal of Experimental Psychology: General*, *130*, 380–400. doi: 10.1037/0096-3445.130.3.380

Tippmann, S. (2015, January). Programming tools: Adventures with R. *Nature*, 109-110. doi:10.1038/517109a

Trafimow, D. (2014). Editorial. *Basic and Applied Social Psychology*, *36*, 1-2. doi: 10.1080/01973533.2014.865505

Trafimow, D. & Marks, M. (2015). Editorial. *Basic and Applied Social Psychology*, *37*, 1-2. doi: 10.1080/01973533.2015.1012991

Vaux, D. (2012, December). Know when your numbers are significant. *Nature*, *492*, 180-181. doi:10.1038/492180a

Weissgerber, T. L., Milic, N. M., Winham, S. J., & Garovic, V. D. (2015). Beyond bar and line graphs: Time for a new data presentation paradigm. *PLOS Biology*, *13*, 1-10. doi: 10.1371/journal.pbio.100212